MDPI

Article

QGenClass: a command line tool that produces classification rules in human readable form

Ioannis G. Tsoulos[1,*], Vasileios Charilogis[2]

[1] Department of Informatics and Telecommunications, University of Ioannina, Greece;itsoulos@uoi.gr
[2] Department of Informatics and Telecommunications, University of Ioannina, Greece; v.charilog@uoi.gr
* Correspondence: itsoulos@uoi.gr

Abstract

Data classification is a computational problem with important applications in many scientific fields and a series of machine learning models have been proposed for effective classification, however most of the time the process either gets trapped in local minima of the error function or produces results that are not interpretable. This paper presents a computational environment that can be used in such problems and which produces understandable classification rules in a high-level programming language. To produce these classification rules, the technique of Grammatical Evolution is utilized. The implemented software includes a series of innovative techniques, such as different initialization techniques for the chromosomes of Grammatical Evolution, incorporation of local optimization rules on the chromosomes, a variety of fitness calculation techniques for the chromosomes, training data improvement techniques for the case of imbalanced data as well as stochastic techniques for terminating the evolutionary process. The implemented software was applied to a wide range of classification problems and the experimental results were compared with other machine learning techniques and the results of these experiments are presented in detail in the present work.

Keywords: Grammatical Evolution; Machine learning; Evolutionary process; Genetic algorithms

Received:
Revised:
Accepted:
Published: